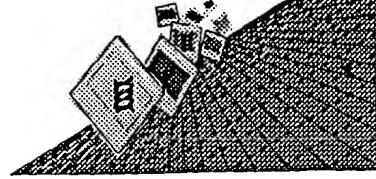


TECHNOLOGY  
SYSTEMS  
BRANCH

DS 90  
09/9

## RAW SEQUENCE LISTING ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/887941

Source: OIPE

Date Processed by STIC: 09/21/01

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

### Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:  
<http://www.uspto.gov/web/offices/pac/checker>

Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	<u>SERIAL NUMBER:</u> <u>09/887947</u>
<b>ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHIA" HEADERS, WHICH WERE INSERTED BY PT</b>		
1 <input type="checkbox"/> Wrapped Nucleic Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 <input type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 <input type="checkbox"/> Variable Length.	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (ii) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
9 <input checked="" type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Sequence(s) _____ missing. If Intentional, please insert the following lines for each skipped sequence: <210> sequence id number <400> sequence id number 000	
10 <input type="checkbox"/> Invalid <213> Response	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents	
11 <input type="checkbox"/> Use of <220>	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or Artificial Sequence	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
13 <input type="checkbox"/> Misuse of n	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.	

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/887,941

DATE: 09/21/2001  
TIME: 20:52:01

Input Set : A:\sequence.txt  
Output Set: N:\CRF3\09212001\I887941.raw

3 <110> APPLICANT: Gonzalgo, Mark L.  
4 Jones, Peter A.  
5 Liang, Gangning  
7 <120> TITLE OF INVENTION: A Cancer Diagnostic Method Based Upon DNA Methylation Differences  
W--> 8 <130> FILE REFERENCE: 47675-21  
W--> 9 <140> CURRENT APPLICATION NUMBER: US 09/887,941  
10 <141> CURRENT FILING DATE: 2001-06-22  
11 <150> PRIOR APPLICATION NUMBER: US 09/094,207  
12 <151> PRIOR FILING DATE: 1998-06-09  
W--> 13 <160> NUMBER OF SEQ ID: 17  
15 <170> SOFTWARE: MS Word

Does Not Comply  
Corrected Diskette Needed

## ERRORED SEQUENCES

127 <210> SEQ ID NO: 9  
128 <211> LENGTH: 145  
129 <212> TYPE: DNA  
130 <213> ORGANISM: Homo sapiens  
132 <400> SEQUENCE: 9  
133 tgagagcagc atcctccctt gcggtgtggtt ctcttaactta cctcctgtat 50  
134 ggggtctgcg gaccagcac acctcccgaa ccccaaaaaa attccagctc 100  
E--> 135 aagagcccta aaaatcccta ccctgnnaaa gtttgagctt ctccc 145

Does Not Comply  
Corrected Diskette Needed

Errored: unknown nucleotides ~~n~~ must  
be described in fields 221, 222 and 223;  
Description should include "unseen", location  
and ~~values~~ possible values of n respectively.

Does Not Comply  
Corrected Diskette Needed